Supplementary

Table S1 Risk of bias assessment of included studies using ROBINS-I

Study (author, year)	Confounding	Selection of participants	Classification of interventions	Deviations from intended interventions	Missing data	Measurement of outcomes	Selection of reported results	Overall judgment
Glenisson *et al.*, 2024	Moderate: multivariate analysis performed; residual confounding may persist	Moderate: single-center; included only patients requiring resection	Low: exposure measurement was based on clinical criteria	Low: no evidence of deviation from planned interventions	Low: follow-up data complete	Low: outcomes are objective and clearly defined	Low: pre-specified outcomes reported	Moderate: risk of bias
Kelm *et al.*, 2021	Serious: groups non-randomized; baseline differences not fully controlled	Moderate: multicenter cohort, but selection was not random	Low: exposure classification is appropriate	Low: no deviations observed	Low: missing data minimal	Serious: some outcomes measured retrospectively	Low: methods reported as planned	Serious: risk of bias
Hammoudi *et al.*, 2020	Moderate: prospective study with multivariate adjustments	Moderate: single-center; selection based on clinical criteria	Low: intervention classification is reliable	Low: no deviations reported	Low: missing data not significant	Low: objective outcome measures	Low: multiple predefined outcomes reported	Moderate: risk of bias
Stevens *et al.*, 2020	Moderate: initial randomization controls for baseline factors; however, the observational, long-term nature (median 5 years) introduces the potential for confounding by subsequent uncontrolled treatments	Low: the cohort was well-defined, comprising all patients from the original LIR!C trial (which was a multicenter RCT)	Low: the initial interventions (resection *vs.* infliximab) were clearly defined and recorded in the original trial	Moderate: patients received additional treatments for Crohn's disease over the 5+ years, which complicates the causal attribution of the initial effect, despite the study reporting these deviations	Low: long-term follow-up was achieved, with assumed low loss to follow-up (standard for high-quality trials)	Moderate: the study is explicitly described as a retrospective study. Although many outcomes are objective (e.g., additional treatment), retrospective collection over an extended period may introduce measurement inconsistencies	Low: the assessed long-term outcomes are a logical and pre-specified extension of the original LIR!C trial's main objective	Moderate: risk of bias
Fortinsky *et al.*, 2017	Moderate: multivariate analysis included important clinical confounders	Moderate: convenience sample; not randomized	Low: exposure clearly classified	Low: no evidence of deviations	Low: missing data low	Moderate: outcome scoring is partly retrospective	Low: predefined outcomes reported	Moderate: risk of bias
Riss *et al.*, 2014	Moderate: risk factor analysis adjusted for key predictors	Low: cohort well defined and consecutively included	Low: no misclassification expected	Low: no deviations from intervention	Low: missing data minimal	Low: surgical recurrence is objective	Low: primary outcome clearly reported	Moderate: risk of bias
Ribbing Wilén *et al.*, 2016	Moderate: retrospective observational design; lack of randomization and limited control of baseline clinical and inflammatory confounders	Moderate: single-center cohort including patients undergoing first ileocolic resection; selection based on clinical indication	Low: surgical intervention (first ileocolic resection) clearly defined, with no misclassification	Low: no relevant deviations from the intended treatment reported	Low: fixed clinical series with no indication of significant loss to follow-up	Moderate: clinical and inflammatory outcomes partly dependent on retrospective records, with potential measurement variability	Low: outcomes consistent with the study objectives and adequately reported	Moderate: risk of bias

https://dx.doi.org/10.21037/tgh-25-131

Table S2 RoB2 risk of bias assessment for Ponsioen *et al*., 2017

RoB2 domain	Judgment	Support for judgment
Bias arising from the randomization process	Some concerns	Randomization via biased-coin minimisation with limited description of sequence concealment; baseline imbalance in smoking status
Bias due to deviations from intended interventions	High risk	Open-label trial with substantial cross-over. Subjective outcomes may be influenced by a lack of blinding
Bias due to missing outcome data	Low risk	Multiple imputation applied; missing data balanced and unlikely to influence outcomes
Bias in the measurement of the outcome	Some concerns	Outcomes (IBDQ, SF-36) are subjective and were assessed without blinding, increasing susceptibility to expectation bias
Bias in the selection of the reported result	Low risk	Outcomes and analyses consistent with pre-registered protocol (NTR1150); no selective reporting identified
Overall RoB2 judgment	High risk of bias	Driven mainly by high risk in deviations from intended interventions and subjective outcomes in an unblinded design

https://dx.doi.org/10.21037/tgh-25-131